<110> Happe, Thomas

<120> Hydrogen Production

<130> 01MEL1

<140> 10/077699
<141> 2002-02-15

<160> 9

<170> PatentIn version 3.2

<210> 1
<211> 5001
<212> DNA
<213> Scenedesmus obliquus

<400> 1
```
gttgctgttg ctgttctaga acaatccata cacacgatta gattgagctc accttcagct      60

cacggaaaat tcttcaggcc tcaacccttc agctccaccc tgcctttctg gaaaaatgca     120

ctcgtggctc tacagggtga gcaaccaggg gcgcaactgc agggcatgct catacagaac     180

atgctgccgc agctgatcat cgctcagcag tgcagtcaag ctgcgcactg gcagcttgca     240

ttgtagctgg tgtacaacat tccagaagcc gactggtatt cgttgcaatt gtcacaattg     300

tgacgcccat gcaaggccca cgagcaatat cgactgcaga accctgtgct gggatctacg     360

ggaatgattg gattggacga tgtcagggcg ttcgacagca ccgtaccaaa gcttgccaaa     420

ctttagcagc ggctgctagc aaccacgaga taagccatgg ccacaacctt gcaacatcgc     480

gcatctgcag ccgccgatgc atgcaaggtc ggtgttgtgc ggttcctgct tgctctgctt     540

caggcaacac agcctccagg tgttcaactt gaaggtgtga caccactggt gtgctggcag     600

ctggccattc ggtttaagcc aagcagtaca gcgctgtcag cttcatcccc gcctggttac     660

tgtgatgtat gtgcttctga tcaagcggtc ctccatgccg tccgaacaga actgcgctgt     720

aagcttacgc agccccaacc ggctccgagc agcatgccct taagtggcgg gaaaactgcc     780

agggacggtg taagggcgcc attcagcgct cgatactgta agattgtttt agatgaaaca     840

gaaatacacc tccggagctg cgagtagcga ggtgattttg cataagggat ccacactgtt     900

gtgggcgcac gtccaagaaa tgtttacccg tttcgattga cagcaaaaca tcatgatcat     960

caaaggagtg catcgacagt caacgatcac caggtgatta cgtttgtcac tgacaagcgc    1020

cctctacgtg cgccttgggc ctacatatgc cctgctgtgg agtacccgt gcacaacaga    1080

gcgttagaga tacttcatag ctgcaactag actacctta ccctaacgaa atcaccctag    1140

accgacagtg tcggagtagc tgcgacccaa acgtgatggc gagcggattg cttctcaagc    1200
```

```
agcgctcggt atgcctgagt ggcaaccggg aggtcggtat gctgtttctg tccgcccgcc     1260

agtgaacagg cgggctgtgg tggcagcagg tgcgcttctt ctgaagggca gctagggctg     1320

tttcgggcag tgcatgccgg cctattttgg gttgctcgga gcaataatat gtactatatt     1380

gctctcgtgg agctgtgttg cgccacgtgc ttgccttggc gcctgttgac cccggaccct     1440

ccacgttgct tcttgccgct gcagagcgca ggcgccttgt tgtgcgggca gctggcccaa     1500

cagcagaatg tgattgccca ccagctcccg cgcccaaggc cccgcactgg cagcagacgc     1560

tagatgagct aggtgagctg cgtgacattg gaagtctggt gtccgcaact gctctctgtg     1620

catgctgaca tccggaatca agtgccaaga agcagggctc gtgtgggtca tttgtgggca     1680

ggtttgcagc agcttgccgt gttcaagcag cagcatgtgg gctgacacat actgctgccg     1740

tgcttctgct gtcctgcagc caagcctaag gagcagcgca aggtgatgat cgcccagatc     1800

gcaccagcag tgcgcgtggc tattgcagag accatgggac tcaaccctgg ggatgtgaca     1860

gttggccaga tggtgaccgg cctgcgcatg ctgggctttg attatgtgtt tggtgagtta     1920

cacagtgttt agtgctgcag cagtccagag cagcttgtgc tagttgatgt tgatcctttg     1980

ggcctgggat atccagctgg acgtcttaca ctgttttttt agcgtccgga gtgggctagt     2040

caacaacagt gagcgctgta tcatgtggtt tgttcatgcg tgcgtcgcat gcatgtggcc     2100

taaccagctg ctgccagcgt gtgcatgtgc ttggtgctgt tttggtgctt ggctggtgag     2160

cagccgcttt ctgtgtttat gtttggctcc tgttccatgc atgttctttg ttctgctgtg     2220

actcatctac tgctgctgct ggtgcatctg ctgcttgcag acacgctgtt tggtgctgac     2280

ctcaccatca tggaggaggg cacagagcta cggcacaggc ttcaggtcag tggtgatggt     2340

gtactgctgt gttcattatg ccatgaggga ctttggtgtt gccatcaaca gctcacactt     2400

gtagttactg gcggtagctg cagcgacagg tggatgcata tcctgcagca catatcctgc     2460

agcaggcagc agcattcatg catgcatccc tttgctcccc tgtctccttg tgctgacagt     2520

gctgcacact agcgccagcc acaccaggga tgtcgataac aatcagtctg atgtcatcca     2580

cggtgtttta aaacacatct cttgctgctt gctgcttgca ggaccacctg gagcagcacc     2640

ccaacaagga ggagccgctg cccatgttca ccagctgctg ccctggctgg gtggccatgg     2700

tggagaagtc caaccccgag ctcatcccct acctgtcttc ctgcaagtcg ccccagatga     2760

tgctgggcgc agtcatcaag aactacttcg ctgccgaggc cggcgccaag cctgaggaca     2820

tctgcaacgt gagcgtgatg ccctgcgtgc gcaagcaggg cgaggctgac cgcgagtggt     2880

tcaacaccac aggggctggc ggcgcgaacg tggaccacgt catgacaact gcagagctgg     2940

gcaagatctt tgtggagcgc ggaatcaagc tgaacgacct gcaggagtcg ccctttgaca     3000
```

```
accccgtcgg cgagggcagc ggcggcggcg tgctgttcgg caccactgga ggcgtgatgg    3060

aggcggcgct gcgcaccgtg tacgaagtgg tgagtgtcag tgtggcggca gctgtgggtt    3120

gtatcgcagc agcagtttgc gcatttggca gtagtgcagc atgtgctggc atgcgcagag    3180

ttgcgcccac ctgtgttgga tgtgagctgg gtttgcatag aggcgccatc tgcagaagcg    3240

tgcacttctg catactgctg ctgctgatct actgccttgc ctttcaccat acccgccacc    3300

cgtaataatc tctcctgctg cactagccct agacagtgcc agacgttgac gctttctgct    3360

gccgtgttgt gcatctccac gcacctctgc tgcaccgcag gtcacacaga agcctttgga    3420

ccgcatcgtc tttgaggacg tgcgcggcct ggagggcatc aaggagtcca cgctgcacct    3480

caccccaggc cccaccagcc ccttcaaggc ctttgcaggc gcagacggca ccggcatcac    3540

cctcaacatc gcggtcgcca acggcctcgg caatgccaag aagctcatca agcagctggc    3600

tgcaggcgag agcaagtacg acttcatcga ggtcatggcc tgccccggcg gctgcatcgg    3660

cggcggcggc cagccgcgca gcgcggacaa gcagatcctg cagaagcgcc aggcggccat    3720

gtacgacctg gacgagcgcg cggtgatccg gcgcagccac gagaacccgc tgattggcgc    3780

gctgtatgag aagttcctgg gcgagcccaa cggccacaag gcgcacgagc tgctgcacac    3840

gcactacgtg gccggcggcg tgcccgatga gaagtgaagc ggtggctggt gatgctggct    3900

gcggcgaaga aacggtgggc atggtggtgg gtgggttgct gcatggtggt gtcgctcgtg    3960

cagcatggtg ggtttgcggt tgtgatgttg ggcatgctgc acggaggtgt ttgcatggtt    4020

atggatatgg ttcaggtgct gtgctgcgtc gcatgccata agcaccttgt gaccctgtgc    4080

gatgcataaa aatagatatt gccatttggt tccaggctgg tggtggcagt ggctggttaa    4140

cagggagtg tgtgtgtttg tgtgtcttca ttgtcggtgt gttcttgctg catgtattgt     4200

agtgtaatgg gttatgcacg cctgcatgcg cacgcgctcc tcgtgctgcg acagtgcaca    4260

acgcacagcg tgatacagct gcaggacgtt gcggaaaaa cacttgttac tggtgacggc     4320

tgaagcagcg atgatggaga gaatggattc gctgctatct cacagggcgt ggctgctgca    4380

tcgccatggc atgtccctgt gcacgcaat gcctgcgta attttgatag tggcagcact      4440

gaggcagctg caaggccttc tgccagcggc tgtttgtgtc ctatctgtgt ttacaggcag    4500

ctgcatttga aggcaagggg gttggccatc actcactttg atcactcact ttgaagcagg    4560

cttccatcca tgtattggtc aacgcactga agttcttttt ttgtcaccag gcagcagtat    4620

tgtgtgcaca ctacttgcta tggagatgac agcagcatca atctcaagca tgatgaaagc    4680

gtatgttgta tcagtgcccc attttgcaga ctcttaagag ctttaccttc tcaggggttg    4740

cagcaggtgg tggtcagcca gttgagggag tgtgtggctg ttgtcttgcc accatgtgag    4800
```

```
tattgaaacc accatcctga gctaagtgtt caggcatctt accctcatac cccgctaccc     4860

tgctactggg agtttcgttt cattgtattg gcagccgttt actaattagt aatggcgctt     4920

gagcgaggca tgtcttgata tgtatgcctt aggagagtgt gagctcaact caattctcat     4980

aagtgtaagc cacacaactg g                                                5001
```

```
<210>   2
<211>   5208
<212>   DNA
<213>   Chlamydomonus rheinhardtii

<400>   2
gatgatatgg atcgtcgtct ggtgctcaag ctcatgggtt ttggctggcc gccgacgctg      60

tcccggaagc accagcagca gcaggggcca gggggtcggg tgatgatgtg ggcgcggtgt     120

atggaggtgg caccctgtat gttcatctgg gcgcttaatt gcgttaagcc attcgagccc     180

acttcggagg caagttcgat ttggtggcgt gagatccgcc tcaccccggt tactgcacgt     240

gcaggagtgg tgtgcagcag tagtcggcag ggtgtcccca ggtattgtgg cgtttgtcgc     300

acggtatgcc ggtgcagtgc tcaggtgcgt aaagcggcgc gtcgcggtgt tggtcgcaac     360

cggatgcttg aagccgaaat cgcttcgtcg gcacgtgtaa gtcattgttt gtttgggtga     420

gctttgctgg cagcgtagaa ccgctgtggc ggacacacgc tcagcaaggg ccaggggggg     480

cgtccaagcc aaggtccaag cgcgcatccc ctcacccctg caccaatgtc caacaccgac     540

agtaatccac gctccgctac gtcgcaagca ggcaatcatg cgtgtctaac atgactgaac     600

tgccccgctg cccgtgaagg gcgatcgtca cgaagttttg tttgcatggt cgtatcggtg     660

tatatgcgca cgcattgttg ccgacacaac ggacacaact atccggctgc ctactgttgt     720

ataagggtca tagaatctag cgttatcctt ccacgagcgt gtggcagcct gctggcgtgg     780

acgagctgtc atgcgttgtt ccgttatgtg tcgtcaaacg ccttcgagcg ctgcccggaa     840

caatgcgtac tagtatagga gccatgaggc aagtgaacag aagcgggctg actggtcaag     900

gcgcacgata gggctgacga gcgtgctgac ggggtgtacc gccgagtgtc cgctgcattc     960

ccgccggatt gggaaatcgc gatggtcgcg cataggcaag ctcgcaaatg ctgtcagctt    1020

atcttacatg aacacacaaa cactctcgca ggcactagcc tcaaaccctc gaaacctttt    1080

tccaacagtt tacaccccaa ttcggacgcc gctccaagct cgctccgttg ctccttcatc    1140

gcaccaccta ttatttctaa tatcgtagac gcgacaagat gtcggcgctc gtgctgaagc    1200

cctgcgcggc cgtgtctatt cgcggcagct cctgcagggc gcggcaggtc gccccccgcg    1260

ctccgctcgc agccagcacc gtgcgtgtag cccttgcaac acttgaggcg cccgcacgcc    1320
```

```
gcctagggtg aggggcgacgc agtgaacgca gtttcgatgg gtcactttgt cgctttttgcg     1380

gaagcctccg aaacgtcccg cgaggttcaa acggccccga atgaccacac ccatatggcc     1440

actgggaata ataacgcagc aacgtcgctt gcgcggctgc cgcacccgct gcggaggcgc     1500

cttttgagtca tgtccagcag gcgctcgccg agcttggagc gaacggccga gcgagcgcgc     1560

acgcattgtt gtggtcaagt ctctccactc agtccgaccc cccacacggc gtagggggtct     1620

gaagtccacc aactcctcac acaccccaag gaagggacgt aagcccccct ggctacgctt     1680

tacccagcag ccacagcgac agagcgcccc aacataggct cgagatagaa cgcacctgaa     1740

ctgtgacact tacaatggaa aggaactgcg gatggcctta aagtcaagca ttttgtgacg     1800

agtcggctcg gaatccccat cggcgcccgt ccgttcgtct tcatcaccgc ctgaaacggc     1860

gcacgcgcaa tagtgcgcac ttgatgcctt tcggtccaac gcctctgtca gctaacactt     1920

tccagggcca gcgcggactc gagaaccctc tttcctggca accttggttt ggctggacct     1980

ggcaaccttg gtttggctgg caccaacctt gacccacata aatctctccc ccccccctt     2040

atgcccacag ccaagcccaa ggacgacccc acgcgcaagc acgtctgcgt gcaggtggct     2100

ccggccgttc gtgtcgctat tgccgagacc ctgggcctgg cgccgggcgc caccacccccc     2160

aagcagctgg ccgagggcct ccgccgcctc ggctttgacg aggtaggtgc gctcgctgct     2220

gcagtgccca acacgcatct tccagctcac cgcctccagt cagcaccttg gcatgcatgc     2280

ttggcgcatc tgccgcctca ttgccgcctc gcggcctcgc cgctgcctgc atcaagcctg     2340

cctgcctgcc tgcccgccct cacgccccag gtgtttgaca cgctgtttgg cgccgacctg     2400

accatcatgg aggagggcag cgagctgctg caccgcctca ccgagcacct ggaggcccac     2460

ccgcactccg acgagccgct gcccatgttc accagctgct gccccggctg gatcggtagc     2520

agcgcggcgt gcttgcttag ggccccataa cctgtcttgg gccccggcg tccgcctctc     2580

cacctacctg caacatgtac gtgcctacgg tattgtcgca tgtctcttga cgatttgggt     2640

cgaccttacc tttgccttgt gtcctttctc cacccccacc cgcctctttc ctcgccggcc     2700

cccctcgcgc agctatgctg gagaaatctt acccggacct gatcccctac gtgagcagct     2760

gcaagagccc ccagatgatg ctggcggcca tggtcaagtc ctacctagcg gaaaagaagg     2820

gcatcgcgcc aaaggacatg gtcatggtgt ccatcatgcc ctgcacgcgc aagcagtcgg     2880

aggctgaccg cgactggttc tgtgtggacg ccgaccccac cctgcgccag ctggaccacg     2940

tcatcaccac cgtggagctg ggcaacatct tcaaggagcg cggcatcaac ctggccgagc     3000

tgcccgaggg cgagtgggac aatccaatgg gcgtgggctc gggcgccggc gtgctgttcg     3060

gcaccaccgg cggtgtcatg gaggcggcgc tgcgcacggt gggtctgtga gagccggttg     3120
```

```
attggcccgg cagaacgcat acacttgctg aacctttgat gcgggataag caaggctacc    3180

gatccgcgtc tttttacacc tgtttatcac gtcgctgagc aagctcgtga cacctgcagg    3240

cctatgagct gttcacgggc acgccgctgc cgcgcctgag cctgagcgag gtgcgcggca    3300

tggacggcat caaggagacc aacatcacca tggtgcccgc gcccgggtcc aagtttgagg    3360

agctgctgaa gcaccgcgcc gccgcgcgcg ccgaggccgc cgcgcacggc acccccgggc    3420

cgctggcctg ggacggcggc gcgggcttca ccagcgagga cggcaggggc ggcatcacac    3480

tgcgcgtggc cgtggccaac gggctgggca acgccaagaa gctgatcacc aagatgcagg    3540

ccggcgaggc caagtacgac tttgtggaga tcatggcctg ccccgcgggc tgtgtgggcg    3600

gcggcggcca gccccgctcc accgacaagg ccatcacgca gaagcggcag gcggcgctgt    3660

acaacctgga cgagaagtga gcgggcggcg ctgctgggat tgggcagggg agggaaggga    3720

ctgcggggca gggtgcggcg ggaaacggaa atgggcaagg ctcgaggtgg agggcggggt    3780

gggttggggt tacttgctac aggttggcgg gcaggatgtg atggaagcag tgtggaggag    3840

gtgtgcgtag ggtcccgacg acggtattcg cacgagcaaa gagggtcggc acttcctgac    3900

acaatgtgcg cctgcacgtg cgctcctgtt gctgccccag gtccacgctg cgccgcagcc    3960

acgagaaccc gtccatccgc gagctgtacg acacgtacct cggagagccg ctgggccaca    4020

aggtgggggg ggttgtatac taccagccca aatgacgggg ctggtcgggg gcgttggaga    4080

ggcgggccgg gagggaggcg ggctgggtgt ggggcaacag caggtgaagg gacggggggg    4140

cgcactgggc agggcggtac atgccttgtc ctgatagcta cccacacgcg actgttgcta    4200

catggatgca tgacgtgtgc cgtgtgcttg acccctgcag gcgcacgagc tgctgcacac    4260

ccactacgtg gccggcggcg tggaggagaa ggacgagaag aagtgaggag cgccagaggc    4320

tctttgggcg gagacagctt caaagcgagg gggcgtatta gcagtaccgt aaatatgcac    4380

tgatgggtga tgcgggtgtc ctcctttata ttgaatgggg tcaaaatagg cggcgggtca    4440

aatgtttcct ttttgagtgg tgtcacagca tggggcacgt gtgcggaggc cagttgccct    4500

ccagtgcacg cgctcccggt gtgtggccgc actggccttg dataatgcac cggtggagga    4560

ttatggaaga gggggactca gaaggctcat tattggacaa tgcctggtct cttccacatt    4620

ggtgtgagcg cggctccgca taggctgttc actgcacgct ggcattaggc gtaggtactg    4680

gcatgaggga cgcggcttg ctaaccgaat ggcgtatccc tccagggcac gtcggaatgg    4740

cgcgtgccca tcaacgcaaa ttcttggcct tcatcgcttc tggatattga agctgcacaa    4800

acctgcattc tatttgcttg tttacacgtg ccccaatctt ggttggaagc taaacatgtt    4860

tgggaacaat tcatcttact aaagcgtgtg ggggttgagg atgcgcacgt tgtgcgctgg    4920
```

```
tgggtgggcg ggaacgtggg tagcatttag ctagctggc atacgacaac ggggcccgtg      4980

aggattgagc acttgactcg cgaacttatg aacgtagcgc tttatacccca ccgtatgcga      5040

ttgacgttgg tgtaggcaac caggcggtag gaaggcggag agatgcattg caaacgcctg      5100

taaaagaacg gcatagctac tagacactct gatgtggacc cttggcgcag ccacgacagg      5160

agaggtgtgc atcagccgct tgtaagcacg cacttctgag aaaaaaaa               5208


<210>  3
<211>  3265
<212>  DNA
<213>  Chlorella fusca

<400>  3
gcggaattac tagtgataag cagtggtaac aacgcagagt cgcgggcagg gactcgatca       60

gttgttatgt gttgccccgt ggttgcaagt aggcacgcag ggcgtgcaag gcatgttgct      120

gtccgtgcag cagggccaac atctgagtgt gattgtcctc caacacctca ggccaagctg      180

cctcactggc agcaggctct ggatgagctc gccaagccca aggagagcag gaggttgatg      240

atcgcgcaaa tcgcctccgc tgttcgtgtc gctattgctg agaccattgg cttggcccca      300

ggagatgtca ccattgggca gctcgtgact gggctgcgta tgcttggctt tgattatgtc      360

tttggtaagc agcagcatct tgcattacac ttgcagttgg tcgtcacatg cacctaatca      420

gatgttagcc ctctggaaca tttttgcctg tttggtgctt acctgaccaa ctgctgcctg      480

gtatggccaa cttgtgaagc tgcgtgtgtt ggcgttgcta cagacaccct gtttggtgct      540

gacctgacca ttatggagga gggaacggag ctgctgcatc gcctgcagga ccatctggag      600

cagcacccca acaaggaggt gagtaagcca gctgggtggt ctaccaccca gcaccagctc      660

gagacagcag ccttgcatca acactcacaa cgtctagctc ctccttaaat gagcggacca      720

aacctgtgag tggcaccatg tcagctgccc ctcgcaccaa agcacagcat ggcctgtctg      780

tcgtcgattg ccacatgagt gtttgcgttg ttatgcaagt gcctgaacaa actgcatatt      840

cctgtgtctc tctgcgttcg cacaggagcc actgcccatg ttcaccagtt gctgcccagg      900

ctgggttgcc atggttgaaa agagcaatcc tgagctcatc ccctacctgt catcttgcaa      960

gtcgcctcag atgatgcttg gggccgttat caagaactac tatgcacagc aggttggagt     1020

gcagcccagt gacatctgca acgtgtcagt catgccatgc gtacgcaagc agggagaggc     1080

tgaccgggag tggttcaaca ccacaggtgg cgcaggcag tgtatcacca gtactggtgt     1140

tctccgtgtg ttgtcagtgt gtctgttaga ggctggatac tctccagtgc agtgctgatg     1200

cagagtggcg gctggtgtgc agcagcgacc ccaagaacac tgagagctgg caattcaatg     1260

ggcttgcttg cttactgtca gcttcctttt cctgcaggtg cagtgacata cggtctgcat     1320
```

```
caaggctcaa acatgttgtg tatgtatgtg tgatgttgca attgcaggcc ttgcccgtga      1380

tgttgatcat gtggtgacta ctgctgaggt tggtaagata ttcctggagc gtggcatcaa      1440

gctgaatgag ctgccagaga gcaactttga caaccccatt ggcgagggca caggtggtgc      1500

tctgctgttt ggcaccactg gaggtgtcat ggaggcagca cttcgcacag tctatgaagt      1560

ggtgagtggt actgcttcag tttcagtcag tgtaccaacc aagctactgc aattgcatag      1620

cgccagtttt ctgccatcaa tgacctgctt tgtaagtagc tgatacttta ccaaccactg      1680

gtatttgtgg ttatcctgcc atagcacatg ccttctcctg ctgttggctt tatcaacctg      1740

ttggtctatg tgtcactgct gtgctgcagg ttacccagaa gcccatgggt cgtgttgact      1800

ttgaggaggt gcgaggcctt gaaggaatca aggaggcaga gatcacactc aagccaggag      1860

acgacagccc attcaaagcc ttcgcaggag ctgatgggca gggcatcacg ctcaagattg      1920

cagtagccaa tgggcttggc aatgccaaga agctcatcaa gagcctgtca gagggcaagg      1980

ccaagtatga tttcattgag gtcatggcat gccctggtgg ctgcattggc ggaggcggtc      2040

agccccgcag tactgacaag cagatcctgc agaagcgcca gcaggctatg tacaacctgg      2100

atgagcgcag taccatccgc cgcagccatg ataacccatt catccaggcg ctgtatgaca      2160

agttcctagg cgcacccaac agccacaagg cacatgatct gctgcacaca cactatgtgg      2220

caggtggaat tccagaggag aagtgaggga ccgaggccgg agtggtgtta ttagtgtaga      2280

gctaggcagc agggatctgg ccgcatttgg gtgctgttgt ttggtttggc atcaaagata      2340

tgatgaatgt acaatctatt gggttctttg tatctcattc atgactgctg cttggtgagg      2400

tatgggccag gaagaagccc gcatcaatgc atgtgaacta ggtggctcca catatgaacc      2460

ctatctggat gtttaaggta cctgaaacaa tagtgcatcg gctctgcatg gctcaacaac      2520

ctgtcttcag agcaggtgta ttccacacca tcttgattta cctaccactc tgtagttcaa      2580

gtggtcaaat tgaatgtcta tggcagctac gcctgcagtt catagtctat gaaggtttca      2640

ccagagtcca tgtccctcat attttttgtt ttatatgcct tgattatgcc ccttgaacca      2700

tgctcaatgc acacaagttg gtcgcaggac aggcggcatc gtacatctca attttcagaa      2760

cttgtcagtg cggcattgcc ttatttgtac tcttgcagtc ctgtttcacc cttgctactg      2820

ccttgcatgc atcttgtttt tgcaagcaac agctcatgca ttgcaatcga tcatcacgta      2880

catccgtgcc atattcacat ggttttgact tgcaaatcaa ccaccaggca gtgggtaaat      2940

tgccaggctg ggtgcacttt gggccatttg ggcagccctc ttgtggcgag ctttgctgca      3000

gggccaagct gagtgcatca gactcagcag gctgctgctg gcactgtaga atgctgaaaa      3060

gggcattcaa ctacatgtca ttattaggtt gacctgagac agccgtaaga atatcattgt      3120
```

```
gtgctgaact tagtcgtcaa tgtcatgcca tgatgtgtgt ttcagggatg gataagggag    3180

gtccttcctc aattacatgc ctttcaagag acttcaatat ctgttgtcag tgacttgttt    3240

gtgtttgctt aatccagtgg ttctc                                         3265
```

<210> 4
<211> 448
<212> PRT
<213> Scenedesmus obliquus

<400> 4

Met Pro Glu Trp Gln Pro Gly Gly Arg Tyr Ala Val Ser Val Arg Pro
1   5   10   15

Pro Val Asn Arg Arg Ala Val Val Ala Ala Glu Arg Arg Arg Leu Val
            20   25   30

Val Arg Ala Ala Gly Pro Thr Ala Glu Cys Asp Cys Pro Pro Ala Pro
            35   40   45

Ala Pro Lys Ala Pro His Trp Gln Gln Thr Leu Asp Glu Leu Ala Lys
        50   55   60

Pro Lys Glu Gln Arg Lys Val Met Ile Ala Gln Ile Ala Pro Ala Val
65   70   75   80

Arg Val Ala Ile Ala Glu Thr Met Gly Leu Asn Pro Gly Asp Val Thr
                85   90   95

Val Gly Gln Met Val Thr Gly Leu Arg Met Leu Gly Phe Asp Tyr Val
                100   105   110

Phe Asp Thr Leu Phe Gly Ala Asp Leu Thr Ile Met Glu Glu Gly Thr
            115   120   125

Glu Leu Leu His Arg Leu Gln Asp His Leu Glu Gln His Pro Asn Lys
    130   135   140

Glu Glu Pro Leu Pro Met Phe Thr Ser Cys Cys Pro Gly Trp Val Ala
145   150   155   160

Met Val Glu Lys Ser Asn Pro Glu Leu Ile Pro Tyr Leu Ser Ser Cys
                165   170   175

Lys Ser Pro Gln Met Met Leu Gly Ala Val Ile Lys Asn Tyr Phe Ala
```

Ala Glu Ala Gly Ala Lys Pro Glu Asp Ile Cys Asn Val Ser Val Met
        195              200              205

Pro Cys Val Arg Lys Gln Gly Glu Ala Asp Arg Glu Trp Phe Asn Thr
    210              215              220

Thr Gly Ala Gly Gly Ala Asn Val Asp His Val Met Thr Thr Ala Glu
225              230              235              240

Leu Gly Lys Ile Phe Val Glu Arg Gly Ile Lys Leu Asn Asp Leu Gln
            245              250              255

Glu Thr Pro Phe Asp Asn Pro Val Gly Glu Gly Ser Gly Gly Val Leu
        260              265              270

Phe Gly Thr Thr Gly Gly Val Met Glu Ala Ala Leu Arg Thr Val Tyr
        275              280              285

Glu Val Val Thr Gln Lys Pro Leu Asp Arg Ile Val Phe Glu Asp Val
    290              295              300

Arg Gly Leu Glu Gly Ile Lys Glu Ser Thr Leu His Leu Thr Pro Gly
305              310              315              320

Pro Thr Ser Pro Phe Lys Ala Phe Ala Gly Ala Asp Gly Thr Gly Ile
            325              330              335

Thr Leu Asn Ile Ala Val Ala Asn Gly Leu Gly Asn Ala Lys Lys Leu
            340              345              350

Ile Lys Gln Leu Ala Ala Gly Glu Ser Lys Tyr Asp Phe Ile Glu Val
        355              360              365

Met Ala Cys Pro Gly Gly Cys Ile Gly Gly Gly Gln Pro Arg Ser
    370              375              380

Ala Asp Lys Gln Ile Leu Gln Lys Arg Gln Ala Ala Met Tyr Asp Leu
385              390              395              400

Asp Glu Arg Ala Val Ile Arg Arg Ser His Glu Asn Pro Leu Ile Gly
            405              410              415

Ala Leu Tyr Glu Lys Phe Leu Gly Glu Pro Asn Gly His Lys Ala His

Glu Leu Leu His Thr His Tyr Val Ala Gly Gly Val Pro Asp Glu Lys
        435                 440                 445


<210>   5
<211>   497
<212>   PRT
<213>   Chlamydomonas rheinhardtii

<400>   5

Met Ser Ala Leu Val Leu Lys Pro Cys Ala Ala Val Ser Ile Arg Gly
1               5                   10                  15


Ser Ser Cys Arg Ala Arg Gln Val Ala Pro Arg Ala Pro Leu Ala Ala
        20                  25                  30


Ser Thr Val Arg Val Ala Leu Ala Thr Leu Glu Ala Pro Ala Arg Arg
        35                  40                  45


Leu Gly Asn Val Ala Cys Ala Ala Ala Ala Pro Ala Ala Glu Ala Pro
    50                  55                  60


Leu Ser His Val Gln Gln Ala Leu Ala Glu Leu Ala Lys Pro Lys Asp
65                  70                  75                  80


Asp Pro Thr Arg Lys His Val Cys Val Gln Val Ala Pro Ala Val Arg
            85                  90                  95


Val Ala Ile Ala Glu Thr Leu Gly Leu Ala Pro Gly Ala Thr Thr Pro
            100                 105                 110


Lys Gln Leu Ala Glu Gly Leu Arg Arg Leu Gly Phe Asp Glu Val Phe
        115                 120                 125


Asp Thr Leu Phe Gly Ala Asp Leu Thr Ile Met Glu Glu Gly Ser Glu
        130                 135                 140


Leu Leu His Arg Leu Thr Glu His Leu Glu Ala His Pro His Ser Asp
145                 150                 155                 160


Glu Pro Leu Pro Met Phe Thr Ser Cys Cys Pro Gly Trp Ile Ala Met
                165                 170                 175


Leu Glu Lys Ser Tyr Pro Asp Leu Ile Pro Tyr Val Ser Ser Cys Lys
        180                 185                 190

```
Ser Pro Gln Met Met Leu Ala Ala Met Val Lys Ser Tyr Leu Ala Glu
        195                 200                 205

Lys Lys Gly Ile Ala Pro Lys Asp Met Val Met Val Ser Ile Met Pro
    210                 215                 220

Cys Thr Arg Lys Gln Ser Glu Ala Asp Arg Asp Trp Phe Cys Val Asp
225                 230                 235                 240

Ala Asp Pro Thr Leu Arg Gln Leu Asp His Val Ile Thr Thr Val Glu
            245                 250                 255

Leu Gly Asn Ile Phe Lys Glu Arg Gly Ile Asn Leu Ala Glu Leu Pro
            260                 265                 270

Glu Gly Glu Trp Asp Asn Pro Met Gly Val Gly Ser Gly Ala Gly Val
        275                 280                 285

Leu Phe Gly Thr Thr Gly Gly Val Met Glu Ala Ala Leu Arg Thr Ala
    290                 295                 300

Tyr Glu Leu Phe Thr Gly Thr Pro Leu Pro Arg Leu Ser Leu Ser Glu
305                 310                 315                 320

Val Arg Gly Met Asp Gly Ile Lys Glu Thr Asn Ile Thr Met Val Pro
            325                 330                 335

Ala Pro Gly Ser Lys Phe Glu Glu Leu Leu Lys His Arg Ala Ala Ala
            340                 345                 350

Arg Ala Glu Ala Ala Ala His Gly Thr Pro Gly Pro Leu Ala Trp Asp
        355                 360                 365

Gly Gly Ala Gly Phe Thr Ser Glu Asp Gly Arg Gly Gly Ile Thr Leu
    370                 375                 380

Arg Val Ala Val Ala Asn Gly Leu Gly Asn Ala Lys Lys Leu Ile Thr
385                 390                 395                 400

Lys Met Gln Ala Gly Glu Ala Lys Tyr Asp Phe Val Glu Ile Met Ala
            405                 410                 415

Cys Pro Ala Gly Cys Val Gly Gly Gly Gly Gln Pro Arg Ser Thr Asp
            420                 425                 430
```

Lys Ala Ile Thr Gln Lys Arg Gln Ala Ala Leu Tyr Asn Leu Asp Glu
         435   440   445

Lys Ser Thr Leu Arg Arg Ser His Glu Asn Pro Ser Ile Arg Glu Leu
         450   455   460

Tyr Asp Thr Tyr Leu Gly Glu Pro Leu Gly His Lys Ala His Glu Leu
465   470   475   480

Leu His Thr His Tyr Val Ala Gly Gly Val Glu Glu Lys Asp Glu Lys
                 485   490   495

Lys

<210> 6
<211> 436
<212> PRT
<213> Chlorella fusca

<400> 6

Met Cys Cys Pro Val Val Ala Ser Arg His Ala Gly Arg Ala Arg His
1   5   10   15

Val Ala Val Arg Ala Ala Gly Pro Thr Ser Glu Cys Asp Cys Pro Pro
                 20   25   30

Thr Pro Gln Ala Lys Leu Pro His Trp Gln Gln Ala Leu Asp Glu Leu
         35   40   45

Ala Lys Pro Lys Glu Ser Arg Arg Leu Met Ile Ala Gln Ile Ala Ser
         50   55   60

Ala Val Arg Val Ala Ile Ala Glu Thr Ile Gly Leu Ala Pro Gly Asp
65   70   75   80

Val Thr Ile Gly Gln Leu Val Thr Gly Leu Arg Met Leu Gly Phe Asp
                 85   90   95

Tyr Val Phe Asp Thr Leu Phe Gly Ala Asp Leu Thr Ile Met Glu Glu
                 100   105   110

Gly Thr Glu Leu Leu His Arg Leu Gln Asp His Leu Glu Gln His Pro
         115   120   125

```
Asn Lys Glu Glu Pro Leu Pro Met Phe Thr Ser Cys Cys Pro Gly Trp
    130             135             140

Val Ala Met Val Glu Lys Ser Asn Pro Glu Leu Ile Pro Tyr Leu Ser
145             150             155             160

Ser Cys Lys Ser Pro Gln Met Met Leu Gly Ala Val Ile Lys Asn Tyr
            165             170             175

Tyr Ala Gln Gln Val Gly Val Gln Pro Ser Asp Ile Cys Asn Val Ser
            180             185             190

Val Met Pro Cys Val Arg Lys Gln Gly Glu Ala Asp Arg Glu Trp Phe
            195             200             205

Asn Thr Thr Gly Ala Gly Leu Ala Arg Asp Val Asp His Val Val Thr
    210             215             220

Thr Ala Glu Val Gly Lys Ile Phe Leu Glu Arg Gly Ile Lys Leu Asn
225             230             235             240

Glu Leu Pro Glu Ser Asn Phe Asp Asn Pro Ile Gly Glu Gly Thr Gly
            245             250             255

Gly Ala Leu Leu Phe Gly Thr Thr Gly Gly Val Met Glu Ala Ala Leu
            260             265             270

Arg Thr Val Tyr Glu Val Val Thr Gln Lys Pro Met Gly Arg Val Asp
            275             280             285

Phe Glu Glu Val Arg Gly Leu Glu Gly Ile Lys Glu Ala Glu Ile Thr
    290             295             300

Leu Lys Pro Gly Asp Asp Ser Pro Phe Lys Ala Phe Ala Gly Ala Asp
305             310             315             320

Gly Gln Gly Ile Thr Leu Lys Ile Ala Val Ala Asn Gly Leu Gly Asn
            325             330             335

Ala Lys Lys Leu Ile Lys Ser Leu Ser Glu Gly Lys Ala Lys Tyr Asp
            340             345             350

Phe Ile Glu Val Met Ala Cys Pro Gly Gly Cys Ile Gly Gly Gly Gly
            355             360             365
```

Gln Pro Arg Ser Thr Asp Lys Gln Ile Leu Gln Lys Arg Gln Gln Ala
   370   375   380

Met Tyr Asn Leu Asp Glu Arg Ser Thr Ile Arg Arg Ser His Asp Asn
385   390   395   400

Pro Phe Ile Gln Ala Leu Tyr Asp Lys Phe Leu Gly Ala Pro Asn Ser
                405   410   415

His Lys Ala His Asp Leu Leu His Thr His Tyr Val Ala Gly Gly Ile
                420   425   430

Pro Glu Glu Lys
        435

<210>   7
<211>   2636
<212>   RNA
<213>   Scenedesmus obliquus

<400>   7
acaacagagc guuagagaua cuucauagcu gcaacuagac uaccuuuacc cuaacgaaau   60

cacccuagac cgacaguguc ggaguagcug cgacccaaac gugauggcga gcggauugcu   120

ucucaagcag cgcucgguau gccugagugg caaccgggag gucgguaugc uguuucuguc   180

cgcccgccag ugaacaggcg ggcuguggug gcagcagagc gcaggcgccu uguugugcgg   240

gcagcuggcc caacagcaga augugauugc ccaccagcuc ccgcgcccaa ggccccgcac   300

uggcagcaga cgcuagauga gcuagccaag ccuaaggagc agcgcaaggu gaugaucgcc   360

cagaucgcac cagcagugcg cguggcuauu gcagagacca uggggacucaa cccugggggau   420

gugacaguug gccagauggu gaccggccug cgcaugcugg gcuuugauua uguguuugac   480

acgcuguuug gugcugaccu caccaucaug gaggagggca cagagcuacu gcacaggcuu   540

caggaccacc uggagcagca ccccaacaag gaggagccgc ugcccauguu caccagcugc   600

ugcccuggcu ggguggccau gguggagaag uccaaccccg agcucaucc cuaccugucu   660

uccugcaagu cgccccagau gaugcugggc gcagucauca agaacuacuu cgcugccgag   720

gccggcgcca agccugagga caucugcaac gugagcguga ugcccugcgu gcgcaagcag   780

ggcgaggcug accgcgagug guucaacacc acaggggcug cggcgcgaa cguggaccac   840

gucaugacaa cugcagagcu gggcaagauc uuuguggagc gcggaaucaa gcugaacgac   900

cugcaggaga cgcccuuuga caaccccguc ggcgagggca gcggcggcgu acuguucggc   960

accacuggag gcgugaugga ggcggcgcug cgcaccgugu acgaaguggu cacacagaag   1020

ccuuuggacc gcaucgucuu ugaggacgug cgcggccugg agggcaucaa ggaguccacg   1080

cugcaccuca ccccaggccc caccagcccc uucaaggccu uugcaggcgc agacggcacc   1140

ggcaucaccc ucaacaucgc ggucgccaac ggccucggca augccaagaa gcucaucaag   1200

cagcuggcug caggcgagag caaguacgac uucaucgagg ucauggccug ccccggcggc   1260

ugcaucggcg gcggcggcca gccgcgcagc gcggacaagc agauccugca gaagcgccag   1320

gcggccaugu acgaccugga cgagcgcgcg ugauccggc gcagccacga gaacccgcug   1380

auuggcgcgc uguaugagaa guuccugggc gagcccaacg gccacaaggc gcacgagcug   1440

cugcacacgc acuacguggc cggcggcgug cccgaugaga agugaagcgg uggcugguga   1500

ugcuggcugc ggcgaagaaa cgguggggcau gguggugggu ggguugcugc auggugguguu   1560

cgcucgugca gcauggugggg uuugcgguug ugauguuggg caugcugcac ggaggguguu   1620

gcaugguuau ggauaugguu caggugcugu gcugcgucgc augccauaag caccuuguga   1680

cccugugcga ugcauaaaaa uagauauugc cauuugguuc caggcuggug guggcagugg   1740

cugguuaaca ggggagugug uguguuugug ugucuucauu gucggugugu ucuugcugca   1800

uguauuguag uguaaugggu uaugcacgcc ugcaugcgca cgcgcuccuc gugcugcgac   1860

agugcacaac gcacagcgug auacagcugc aggacguuug cggaaaaaca cuuguuacug   1920

gugacggcug aagcagcgau gauggagaga auggauucgc ugcuaucuca cagggcgugg   1980

cugcugcauc gccauggcau gucccuguug cacgcaauug ccugcguaau uuugauagug   2040

gcagcacuga ggcagcugca aggccuucug ccagcggcug uuuguguccu aucuguguuu   2100

acaggcagcu gcauuugaag gcaagggggu uggccaucac ucacuuugau cacucacuuu   2160

gaagcaggcu uccauccaug uauuggucaa cgcacugaag uucuuuuuuu gucaccaggc   2220

agcaguauug ugugcacacu acuugcuaug gagaugacag cagcaucaau cucaagcaug   2280

augaaagcgu auguuguauc agugccccau uuugcagacu cuuaagagcu uuaccuucuc   2340

agggguugca gcagguggug gucagccagu ugagggagug uguggcuguu gucuugccac   2400

caugugagua uugaaaccac cauccugagc uaaguguuca ggcaucuuac ccucauaccc   2460

cgcuacccug cuacugggag uuucguuuca uuguauuggc agccguuuac uaauuaguaa   2520

uggcgcuuga gcgaggcaug ucuugauaug uaugccuuag gagaguguga gcucaacuca   2580

auucucauaa guguaagcca cacaacugga aaaaaaaaaa aaaaaaaaaa aaaaaa   2636

<210>   8
<211>   2399
<212>   RNA
<213>   Chlamydomonas rheinhardtii

<400> 8

```
aucuuacaug aacacacaaa cacucucgca ggcacuagcc ucaaacccuc gaaaccuuuu      60

uccaacaguu uacaccccaa uucggacgcc gcuccaagcu cgcuccguug cuccuucauc     120

gcaccaccua uuauuucuaa uaucguagac gcgacaagau gucggcgcuc gugcugaagc     180

ccugcgcggc cgugucuauu cgcggcagcu ccugcagggc gcggcagguc gccccccgcg     240

cuccgcucgc agccagcacc gugcguguag cccuugcaac acuugaggcg cccgcacgcc     300

gccuaggcaa cgucgcuugc gcggcugccg cacccgcugc ggaggcgccu uugagucaug     360

uccagcaggc gcucgccgag cuugccaagc ccaaggacga ccccacgcgc aagcacgucu     420

gcgugcaggu ggcuccggcc guucgugucg cuauugccga gacccugggc cuggcgccgg     480

gcgccaccac ccccaagcag cuggccgagg gccuccgccg ccucggcuuu gacgaggugu     540

uugacacgcu guuuggcgcc gaccugacca ucauggagga gggcagcgag cugcugcacc     600

gccucaccga gcaccuggag gcccacccgc acuccgacga gccgcugccc auguucacca     660

gcugcugccc cggcuggauc gcuaugcugg agaaaucuua cccggaccug aucccctacg     720

ugagcagcug caagagcccc cagaugaugc uggcggccau ggucaagucc uaccuagcgg     780

aaaagaaggg caucgcgcca aaggacaugg ucaugguguc caucaugccc ugcacgcgca     840

agcagucgga ggcugaccgc gacugguucu guguggacgc cgaccccacc cugcgccagc     900

uggaccacgu caucaccacc guggagcugg gcaacaucuu caaggagcgc ggcaucaacc     960

uggccgagcu gcccgagggc gagugggaca auccaauggg cguggcucg ggcgccggcg    1020

ugcuguucgg caccaccggc ggugucaugg aggcggcgcu gcgcacggcc uaugagcugu    1080

ucacgggcac gccgcugccg cgccugagcc ugagcgaggu gcgcggcaug gacggcauca    1140

aggagaccaa caucaccaug gugcccgcgc ccggguccaa guuugaggag cugcugaagc    1200

accgcgccgc cgcgcgcgcc gaggccgccg cgcacggcac ccccgggccg cuggccuggg    1260

acggcggcgc gggcuucacc agcgaggacg gcaggggcgg caucacacug cgcguggccg    1320

uggccaacgg gcugggcaac gccaagaagc ugaucaccaa gaugcaggcc ggcgaggcca    1380

aguacgacuu uguggagauc auggccugcc ccgcgggcug uguggcggc ggcggccagc    1440

cccgcuccac cgacaaggcc aucacgcaga agcggcaggc ggcgcuguac aaccuggacg    1500

agaaguccac gcugcgccgc agccacgaga acccguccau ccgcgagcug uacgacacgu    1560

accucggaga gccgcugggc cacaaggcgc acgagcugcu gcacacccac uacguggccg    1620

gcggcgugga ggagaaggac gagaagaagu gaggagcgcc agaggcucuu ugggcggaga    1680

cagcuucaaa gcgagggggc guauuagcag uaccguaaau augcacugau ggguggaugcg    1740
```

```
ggugucucc uuuauauuga auggggucaa aauaggcggc gggucaaaug uuuccuuuuu    1800

gagugguguc acagcauggg gcacgugugc ggaggccagu aggcuguuca cugcacgcug    1860

gcauuaggcg uagguacugg caugagggag cgcggcuugc uaaccgaaug gcguaucccu    1920

ccagggcacg ucggaauggc gcgugcccau caacgcaaau ucuuggccuu caucgcuucu    1980

ggauauugaa gcugcacaaa ccugcauucu auuugcuugu uuacacgugc cccaaucuug    2040

guuggaagcu aaacauguuu gggaacaauu caucuuacua aagcgugugg ggguugagga    2100

ugcgcacguu gugcgcuggu gggugggcgg gaacgugggu agcauuuagg cuagcuggca    2160

uacgacaacg gggcccguga ggauugagca cuugacucgc gaacuuauga acguagcgcu    2220

uuauacccac cguaugcgau ugacguuggu guaggcaacc aggcgguagg aaggcggaga    2280

gaugcauugc aaacgccugu aaaagaacgg cauagcuacu agacacucug auguggaccc    2340

uuggcgcagc cacgacagga gaggugugca ucagccgcuu guaagcacgc acuucugag    2399
```

<210> 9
<211> 2421
<212> RNA
<213> Chlorella fusca

<400> 9
```
gcggaauuac uagugauaag cagugguaac aacgcagagu cgcgggcagg gacucgauca     60

guuguuaugu guugccccgu gguugcaagu aggcacgcag ggcgugcaag gcauguugcu    120

guccgugcag cagggccaac aucugagugu gauuguccuc caacaccuca ggccaagcug    180

ccucacuggc agcaggcucu ggaugagcuc gccaagccca aggagagcag gagguugaug    240

aucgcgcaaa ucgccuccgc uguucguguc gcauugcug agaccauugg cuuggcccca    300

ggagauguca ccauugggca gcucgugacu gggcugcgua ugcuuggcuu ugauuauguc    360

uuugacaccc uguuuggugc ugaccugacc auuauggagg agggaacgga gcugcugcau    420

cgccugcagg accaucugga gcagcacccc aacaaggagg agccacugcc cauguucacc    480

aguugcugcc caggcugggu ugccaugguu gaaaagagca auccugagcu cauccccuac    540

cugucaucuu gcaagucgcc ucagaugaug cuuggggccg uuaucaagaa cuacuaugca    600

cagcagguug gagugcagcc cagugacauc ugcaacgugu cagucaugcc augcguacgc    660

aagcagggag aggcugaccg ggagugguuc aacaccacag gugcaggccu ugcccgugau    720

guugaucaug uggugacuac ugcugagguu gguaagauau uccuggagcg uggcaucaag    780

cugaaugagc ugccagagag caacuuugac aaccccauug gcgagggcac agguggugcu    840

cugcuguuug gcaccacugg agguguauug gaggcagcac uucgcacagu cuaugaagug    900

gugacccaga agcccauggg ucguguugac uuugaggagg ugcgaggccu ugaaggaauc    960
```
```

```
aaggaggcag agaucacacu caagccagga gacgacagcc cauucaaagc cuucgcagga     1020

gcugaugggc agggcaucac gcucaagauu gcaguagcca augggcuugg caaugccaag     1080

aagcucauca agagccuguc agagggcaag gccaaguaug auuucauuga ggucauggca     1140

ugcccuggug gcugcauugg cggaggcggu cagccccgca guacugacaa gcagauccug     1200

cagaagcgcc agcaggcuau guacaaccug gaugagcgca guaccauccg ccgcagccau     1260

gauaacccau ucauccaggc gcuguaugac aaguuccuag gcgcacccaa cagccacaag     1320

gcacaugauc ugcugcacac acacuaugug gcagguggaa uuccagagga gaagugaggg     1380

accgaggccg gagugguguu auuaguguag agcuaggcag cagggaucug gccgcauuug     1440

ggugcuguug uuugguuugg caucaaagau augaugaaug uacaaucuau uggguucuuu     1500

guaucucauu caugacugcu gcuuggugag guaugggcca ggaagaagcc cgcaucaaug     1560

caugugaacu agguggcucc acauaugaac ccuaucugga uguuuaaggu accugaaaca     1620

auagugcauc ggcucugcau ggcucaacaa ccugucuuca gagcaggugu auuccacacc     1680

aucuugauuu accuaccacu cuguaguuca aguggucaaa uugaaugucu auggcagcua     1740

cgccugcagu ucauagucua ugaagguuuc accagagucc auguccuuca uauuuuuugu     1800

uuuauaugcc uugauuaugc cccuugaacc augcucaaug cacacaaguu ggucgcagga     1860

caggcggcau cguacaucuc aauuuucaga acuugucagu gcggcauugc cuuauuugua     1920

cucuugcagu ccuguuucac ccuugcuacu gccuugcaug caucuuguuu uugcaagcaa     1980

cagcucaugc auugcaaucg aucaucacgu acauccgugc cauauucaca ugguuuugac     2040

uugcaaauca accaccaggc aguggguaaa uugccaggcu ggugcacuu ugggccauuu     2100

gggcagcccu cuuguggcga gcuuugcugc agggccaagc ugagugcauc agacucagca     2160

ggcugcugcu ggcacuguag aaugcugaaa agggcauuca acuacauguc auuauuaggu     2220

ugaccugaga cagccguaag aauaucauug ugugcugaac uuagucguca augucaugcc     2280

augaugugug uuucagggau ggauaaggga gguccuuccu caauuacaug ccuucaaga     2340

gacuucaaua ucuguuguca gugacuuguu uguguuugcu uaauccagug guucucaaaa     2400

aaaaaaaaaa aaaaaaaaaa a                                              2421
```